



GB 2291991A

**(12) UK Patent Application (19) GB (11) 2 291 991 (13) A**

(43) Date of A Publication 07.02.1996

<p>(21) Application No 9519670.5</p> <p>(22) Date of Filing 27.09.1995</p>	<p>(51) INT CL<sup>6</sup> G06F 12/12 12/02</p> <p>(52) UK CL (Edition O ) G4A AMG1 AMX</p> <p>(56) Documents Cited GB 2251324 A EP 0522780 A2 WO 94/20906 A1</p> <p>(58) Field of Search UK CL (Edition N ) G4A AMG1 AMX INT CL<sup>6</sup> G06F 3/00 12/00 12/02 12/06 12/08 12/12 On-line : WPI, INSPEC, COMPUTER</p>
<p>(71) Applicant(s) <b>Memory Corporation</b>  (Incorporated in the United Kingdom)  <b>The Computer House, Dalkeith Palace, Dalkeith, EDINBURGH, EH22 2NA, United Kingdom</b></p> <p>(72) Inventor(s) <b>Alan Welsh Sinclair</b></p> <p>(74) Agent and/or Address for Service <b>Rory Macleod</b> <b>The Computer House, Dalkeith Palace, Dalkeith, EDINBURGH, EH22 2NA, United Kingdom</b></p>	

**(54) Disk drive emulation with a block-erasable memory**

(57) A solid state memory for emulating a disk drive comprises means for translating a logical sector address to a main memory address and a main memory composed of non-volatile memory cells erasable in blocks. A write pointer is used to point to an unwritten location in main memory, and an erase pointer is used to point to the next unerased erasable block in address-space sequence to the erasable block containing the unwritten memory location; control means being provided to ensure that there is always at least one erasable block in the erased condition between the two pointers. The memory cells may be of the Flash EPROM, chalcogenide or ferro-optic type.

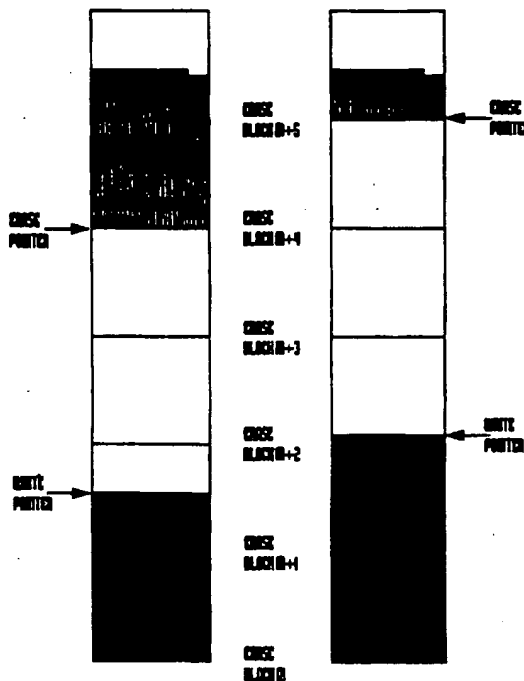


FIGURE 3A

FIGURE 3B

GB 2 291 991 A

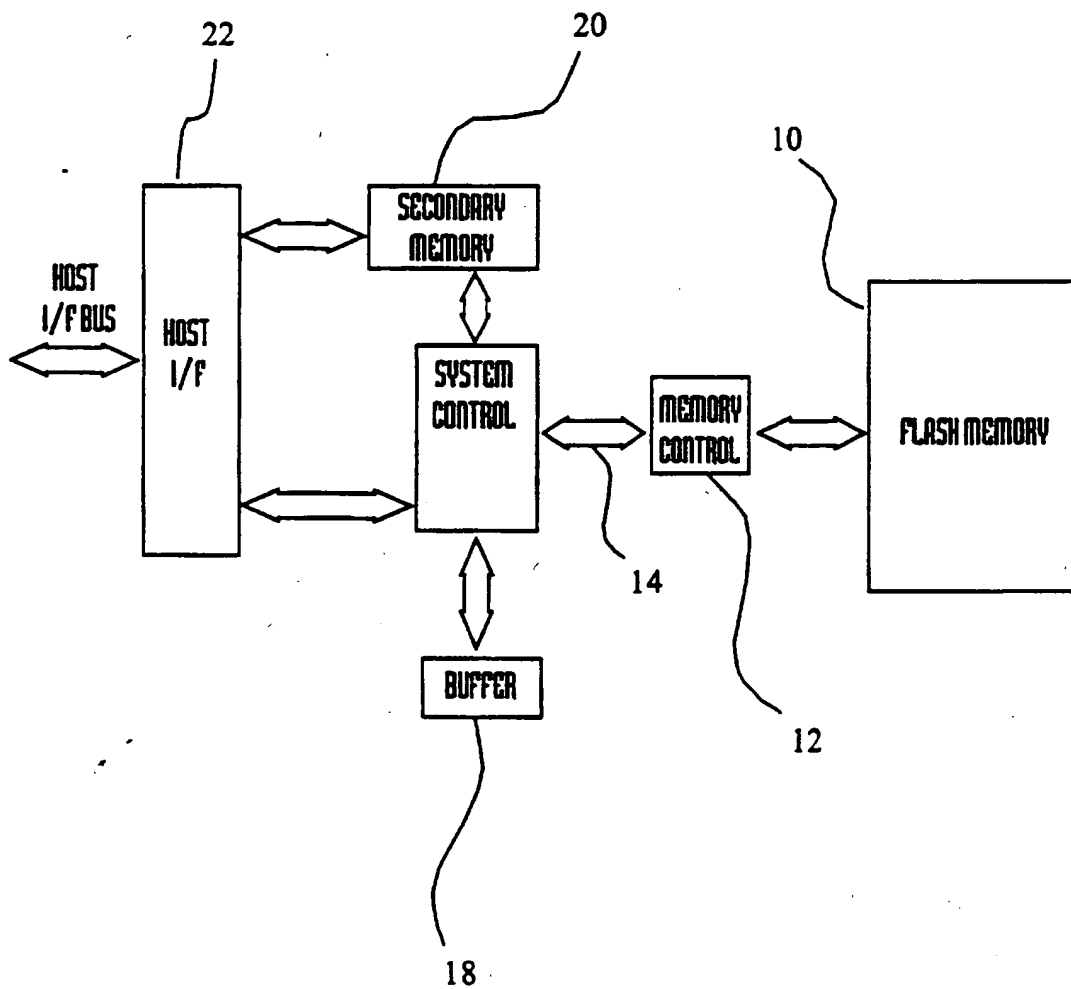


FIGURE 1

2/3

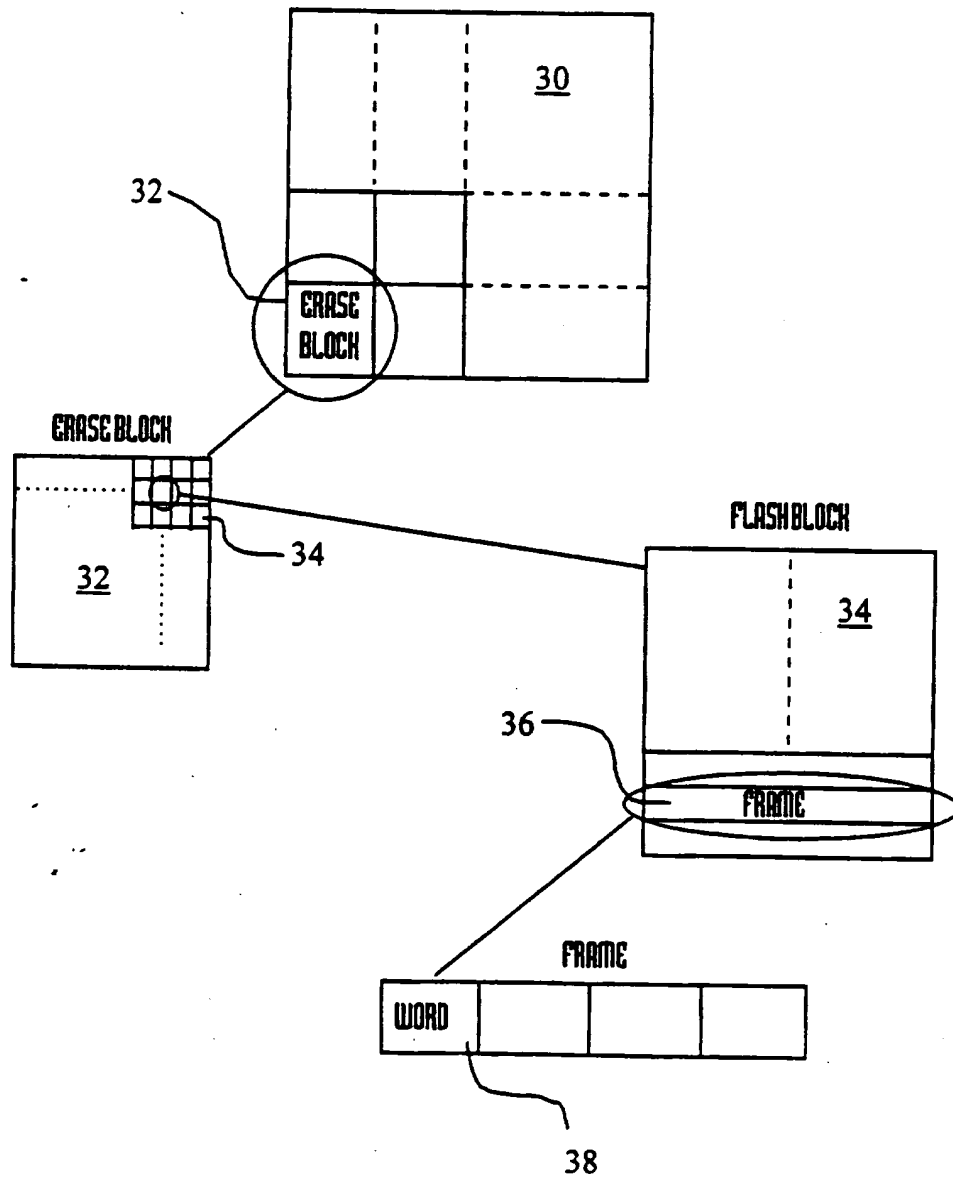


FIGURE 2

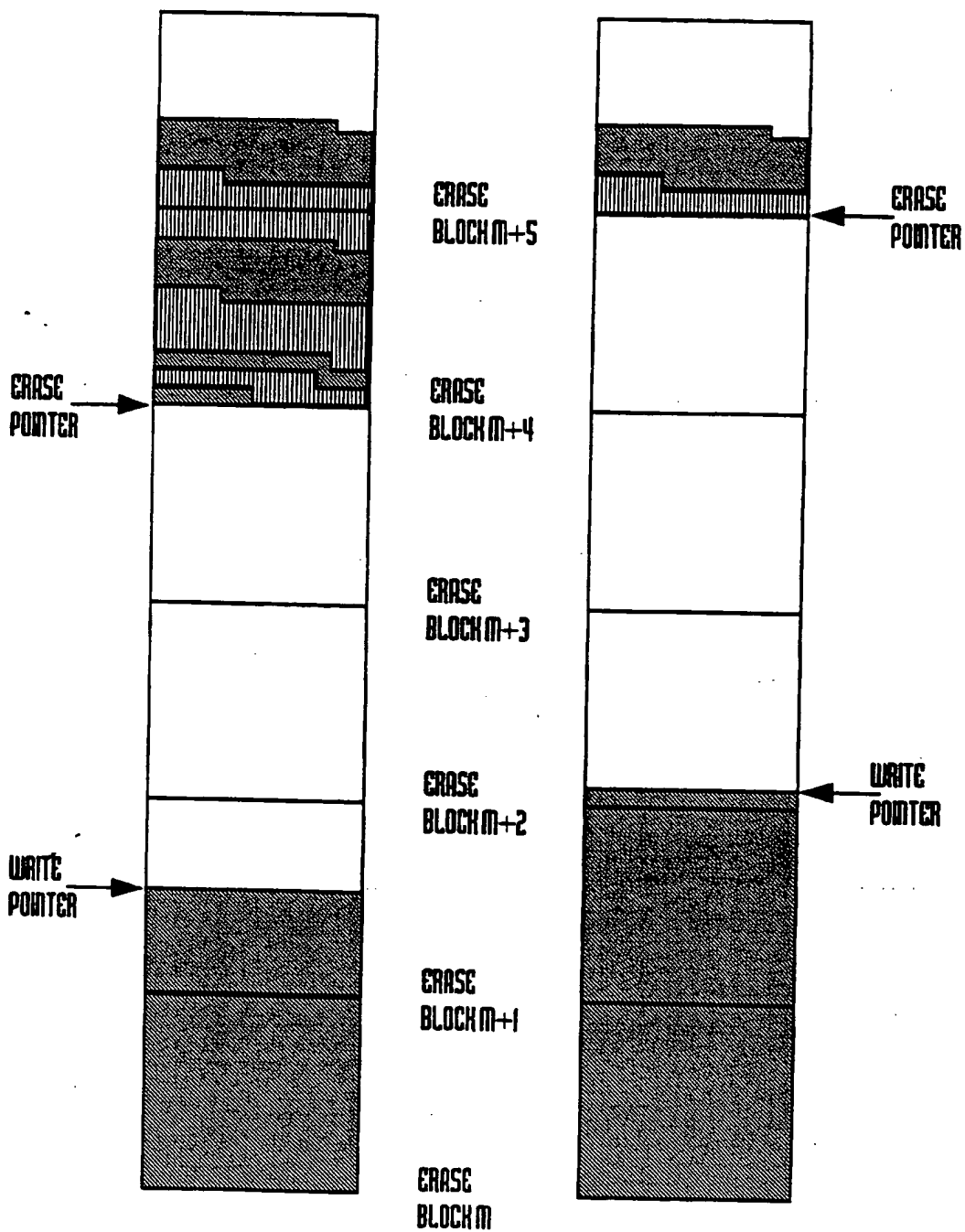


FIGURE 3A

FIGURE 3B

## Algorithm for an Improved FLASH Memory Storage System

The present invention relates to semiconductor memory systems. In particular it relates to memory systems for storing block structured data. Non-volatile memory cells are used as the storage medium. The particular type of non-volatile memory cells used in this invention are memory cells which do not allow erasure of individual cells but allow erasure of blocks of  
5 memory cells. This invention could be used with a number of different memory cells, such as Flash EPROM cells, chalcogenide memory cells, and ferro-optic cells.

Many common memory storage devices for computers are based on rotating storage media such as the magnetic disk drive. Rotating storage media have, however, a few disadvantages. They require comparatively large amounts of power because they consist of rotating platters.  
10 They are comparatively slow to read from and write to because of the seek latency (the time taken to move the head to the correct location) and the rotational latency (the time taken to locate the correct part of a track). Magnetic disks are also comparatively heavy and not very robust (they are sensitive to being knocked). In contrast, semiconductor memory is lightweight, robust, consumes very little power, and is fast in operation. For these reasons it is  
15 desirable to use a solid state equivalent of the magnetic disk. The particular semiconductor chosen as the solid state memory should be high density, non-volatile, and low cost per bit. This makes FLASH memory an ideal candidate.

There are many prior art references to the use of FLASH memory as the storage medium for a solid state disk drive. These references include the following patents: EP 0 392 895,  
20 EP 0 424 191, EP 0 522 780, EP 0 597 706, EP 0 617 363, EP 0 618 535, EP 0 619 541, GB 2 251 323, WO 93 02418, WO 94 19746, WO 94 20906, WO 94 23369, WO 94 23432, WO 95 14272, WO 95 10083, US 4 511 964, WO 84 00628, WO 95 10083, WO 95 14272. Many of these references have certain features in common. For example, most references have a table for converting the logical address from the host computer (usually in CHS - cylinder,  
25 head and sector - format) to a physical address in the FLASH memory or to an intermediate address which will be used for generating the physical address in FLASH memory. Most references also have some means for ensuring that one block is not written to substantially more times than another block. A third common feature in FLASH solid state disk emulators is a buffer for increasing the write speed to the semiconductor. The buffer is usually made of  
30 some fast memory such as SRAM (Static Random Access Memory).

Any solid state disk emulation system which uses FLASH memory as the storage medium

must also have some method of controlling the erasure and writing of the cells. The reason for this is that the cells take a comparatively long time to erase and a cell can only be written to once before it needs to be erased. The present invention is an algorithm which can be used in a solid state disk emulating system for controlling the writing and erasure of memory cells.

- 5 EP 0 522 780 describes a control method for a solid state disk which uses a memory managing table recording the number of times each cell has been erased, the status of each memory block and a buffer to hold data which is to be written to memory. The method involves copying sectors which are to be erased to another block before erasing the initial block.

- EP 0 392 895 relates to the use of a cache memory for buffering writes to the FLASH memory  
10 to increase the speed of operation of the memory. The memory system in that disclosure uses means to determine "the time since each data file was last written". It also relates to correcting for errors and defect substitution, but these resulted in divisional applications which are discussed below.

- EP 0 618 535 relates to a solid state disk emulator which has defect substitution. EP 0 617 363  
15 relates to a method of correcting for defective cells in a FLASH memory by using pointers to point from the defective locations in the memory to the spare locations.

EP 0 424 191 relates to a method of storing sequential data in a memory which is accessible in a prescribed order. Defects in the memory are stored in an ordered directory so that these defective locations are skipped when the memory is being written to.

- 20 EP 0 619 541 provides a solid state disk emulating system which allows random access from a central processing unit. It also measures the erasure count for each block of memory so that a block can be avoided if it has been written to a large number of times.

EP 0 597 706 discloses a solid state peripheral storage device which uses a MAP ROM to map the logical sector address provided by the host to a physical address in the semiconductor

- 25 memory. The addresses of any bad sectors in semiconductor memory are recorded in the MAP ROM. A microsequencer is used to control the mapping.

GB 2 251 323 describes a disk emulation system for FLASH memory. The clean-up operation described in that disclosure is only automatic in the sense that it happens as a background task when certain conditions are met. Either the least cycled block is erased or the

- 30 block with the highest number of deleted sectors is erased. This gives rise to an essentially random block selection for erasure.

WO 94 20906 discloses a FLASH file system which is used to emulate the operation of a disk

drive. It uses a block allocation map to store information on the blocks in FLASH memory that can be written to. When a block is to be updated the allocation map is scanned until a free block is located. A transfer unit is used to facilitate memory reclaiming of the system. The transfer unit is an unwritten erased block of memory. The active (currently valid) data from a block of memory that contains old data (the old block) is written to the same locations in the transfer unit (the new block) and the old block is erased. The old block then becomes the transfer unit.

WO 94 23432 discloses a FLASH memory mass storage architecture. Erase cycles are avoided by programming altered data file into an empty mass storage block. Periodically the storage is cleaned up with a multi-sector erase. A counter is used to ensure that no block is erased more than a maximum number of times.

WO 94 23369 discloses a FLASH memory mass storage architecture which is very similar to that disclosed in WO 94 23432. Both disclosures are by the same applicant.

WO 94 19746 discloses a FLASH solid state drive which stores two bytes at a time. Each time a sector is written into FLASH memory a header is also written with it. If an older version of that sector exists in the memory then it is marked as invalid. When storage space dips below a certain threshold then a clean up operation is performed. Good data from a sector is copied to a buffer, the sector is erased and the good data is copied back to the sector.

WO 84 00628 discloses a method of managing defects in a storage media. The method involves receiving an address from a host, compensating the received address for the number of defective locations below that address, and accessing the compensated address. A sequentially ordered list of defective locations in the storage media is constructed.

WO 95 10083 uses counters to count the number of erase cycles each block has experienced and copies updated versions of data files to new locations. Periodically, blocks are erased to clean up the system. The blocks in the semiconductor memory are sized to conform to commercial hard disk sector sizes.

WO 95 14272 provides a method of generating the correct physical starting address from the number of the logical block to be accessed.

One advantage of the invention is that it provides automatic equalisation of erase/write cycles throughout the entire memory. This automatic equalisation is automatic in the fullest sense in that the equalisation occurs as data is being written, rather than as a background task at some later time. Another advantage of the present invention is that it provides automatic

compaction of the stored data. Yet another advantage of the present invention is that it achieves 100% efficiency in the use of available FLASH memory capacity, that is, all available FLASH memory is used. A further advantage of the present invention is that it can be used with FLASH memory with any erasable block size: the erasable block size does not have to be 5 512 bytes. This enables error correction codes to be included with each segment of data stored. The invention also provides for a very simple translation from physical to FLASH address space.

The present invention provides a solid state memory for emulating a disk drive comprising: translation means for translating a logical address to a second address for accessing a store; a 10 store composed of non-volatile memory cells erasable in blocks; characterised in that a first pointer is used to point to an unwritten location in the store, and a second pointer is used to point to the next unerased erasable block in sequence to the erasable block containing the said unwritten memory location; control means being provided to ensure that there is always at least one erasable block in the erased condition between the first and second pointers.

15 A solid state memory where the said unwritten location in the store is the lowest unwritten location in the store.

A solid state memory where the said unwritten location in the store is the highest unwritten location in the store.

A solid state memory where the said next unerased erasable block in sequence to the erasable 20 block containing the said unwritten memory location is the unerased erasable block with the lowest physical address above the first pointer.

A solid state memory where the said next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is the unerased erasable block with the highest physical address below the first pointer.

25 A solid state memory where the said next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is determined by a suitable algorithm.

A solid state memory where a buffer is used to increase the speed of writing data to the the store.

A solid state memory where the said buffer is an SRAM buffer.

30 A solid state memory where the said buffer is a DRAM buffer.

A solid state memory where the said store is FLASH EPROM.

A solid state memory where the said the store is composed of chalcogenide cells.



A solid state memory where the said store is a ferro-optic memory.

A solid state memory where the said control means are provided by a microprocessor, a microcontroller, or a state machine.

In the present application an unwritten memory location does not refer to a memory location  
 5 which has never been written to, but rather it refers to a location which is in the erased condition.

In the present application a second address may be a physical address, a physical block address, or an intermediate address.

Figure 1 shows a diagram of a memory system for emulating a magnetic disk drive.

10 Figure 2 shows the hierarchy of subdivisions within a FLASH memory.

Figures 3A and 3B illustrate how the write/erase algorithm manages the location and frequency of erase and write operations in main FLASH memory. Figure 3A shows a FLASH memory prior to the erase/write algorithm taking effect, Figure 3B shows the same FLASH memory after the write/erase algorithm has taken effect.

15 Referring to Figure 1, a FLASH memory 10 is used as the main store for data received from the host. The FLASH memory 10 is controlled by a memory controller 12. The memory controller 12 operates between the FLASH memory 10 and a technology independent interface 14. The memory controller 12 controls the low-level execution of high level functions within the physical memory (FLASH memory 10). The memory controller 12 is memory technology  
 20 dependent.

The technology independent interface 14 connects the memory controller 12 to a system controller 16. The system controller 16 manages the storage and accessing of block-structured data in the main FLASH memory 10 and in a write buffer 18 which is included to increase the speed of writing data to FLASH memory 10. The system controller 16 performs the following  
 25 functions. It translates the logical address from the host to an address suitable for use with the FLASH memory 10, hereinafter referred to as the logical to physical address translation. This translation may result in one or more intermediate addresses being generated. It maintains any address lookup tables which are needed. It manages defective storage locations in the FLASH memory 10. It manages the write buffer 18. It manages the block erasure of FLASH memory  
 30 10. It manages the start-up and shut-down of the system after application and removal of power.

All sector writes from the host are directed to the write buffer 18. This minimises the effect of

the slow write speed of FLASH memory. Sector data is transferred from the write buffer 18 to the FLASH memory 10 as a background task. In applications where the write speed is not important the write buffer 18 may be omitted.

The system controller 16 is connected to a secondary memory 20 and to a host interface 22.

- 5 The secondary memory 20 is used to store information such as defective block and erasable block tables for FLASH memory 10, address translation tables, error logs, and microcontroller firmware. The secondary memory may be directly accessible from the host interface 22 to provide a directly mapped memory address space to the host.

Referring to Figure 2, a main FLASH memory 10 is arranged in the following hierarchy of

- 10 subdivisions. A FLASH memory consists of one or more semiconductor chips 30. A semiconductor chip 30 is a physical subdivision which does not have any logical significance. An erase block 32 is a subdivision of a semiconductor chip 30. It is the minimum number of cells that may be deleted in an erase operation. The number of erase blocks in a semiconductor chip varies according to the size of the chip and the chip configuration. A rectangular group of word locations is assigned as a FLASH block. A FLASH block 34 is the minimum maskable block size for defect management. If a cell is faulty then the minimum area which may be masked and replaced is a FLASH block 34. Defect mapping has a granularity of one FLASH block 34. A sector is one data unit according to disk emulation protocols. The data stored in FLASH memory (typically 512 bytes plus ECC bytes and a header) will hereinafter be referred to as a FLASH sector: the data sent by the host operating system (typically 512 bytes) will hereinafter be referred to as a host sector. A host sector is the basic unit of data transfer with the host. The most common host sector size is 512 bytes. A frame 36 is an addressable sub-unit of a FLASH block 34, for example a row within a rectangular FLASH block 34. Each word 38 is randomly accessible as the minimum addressable unit at the technology independent interface 14. The word size is set by the width of the data I/O to the memory system: it may be one, two, or four bytes.

The actual physical location of bits within a word is transparent to the algorithm used in the present invention. The FLASH memory is treated simply as a rectangular array of word locations.

- 30 The FLASH block 34 could be set to equal the host sector size (in most cases the host sector size is 512 bytes), the frame size or even the word size. The memory harvest (the fraction of gross memory that can be used for data storage) reaches a plateau below a certain value of

FLASH block size, so that further reduction of the FLASH block size does not result in much improvement in the memory harvest. There are, however, disadvantages associated with having a small FLASH block size. The smaller the FLASH block size the more storage management space is required. The component test complexity also increases. Thus, there is  
 5 an optimum range for the size of a FLASH block.

Different addressing formats could be used with this invention without departing from the inventive concept.

For 100% memory utilisation the start address for a sector must be the word immediately following the end address of the preceding sector. To accomplish this the present invention  
 10 allows the start address to be defined to a resolution of a single word. Since the FLASH sectors are stored without any intervening gaps, FLASH sectors will overflow both FLASH blocks and erase blocks. In such a case a FLASH sector is known as a split sector and its continuation region is located at the start of the following FLASH block or erase block. The start address of FLASH sectors within FLASH blocks will therefore be different for each  
 15 FLASH block. To ensure that this is easily manageable, it is necessary to ensure that the order in which erase blocks are written corresponds exactly to their physical addressing order. If there is a defective cell in a FLASH block then the block will be masked out. This may give rise to the situation where part of a FLASH sector is stored in one FLASH block, the next FLASH block is masked out, and the rest of the FLASH sector is stored in the FLASH block  
 20 immediately above the one that was masked out. FLASH memory space is therefore discontinuous.

Figures 3A and 3B illustrate how the write/erase algorithm manages the location and frequency of erase and write operations in main FLASH memory. Figure 3A shows a FLASH memory prior to the erase/write algorithm taking effect, Figure 3B shows the same FLASH  
 25 memory after the write/erase algorithm has taken effect.

The algorithm is based on the existence of two pointers, the write pointer and the erase pointer, defining current write and erase locations. The locations are defined as physical sector addresses and the pointers cycle through physical sector address space.

The write pointer defines the physical sector number to which the next sector write operation  
 30 should be directed. The physical location of this sector is guaranteed to have been previously erased in preparation for writing data by the system controller 16 in the memory system.

Sector data to be written may have originated from either the write buffer 18 following a sector

write from the host or from another sector in main FLASH memory 10 in preparation for an erase block erase operation. Following a sector write operation, the write pointer is incremented by one to point to the next highest physical sector address. When the write pointer reaches the highest usable physical sector address it then wraps around and points at the lowest usable physical address.

The erase pointer defines the erase block address in FLASH memory space of the next erase block which is scheduled for erasure. A second field in the erase pointer defines the physical sector number of the first complete sector located at the bottom of the erase block. Following an erase block erase operation, the erase block number is incremented by one. A lookup table (hereinafter referred to as the erase block table) provides the physical sector address of the first sector in the erase block that the erase pointer is pointing at. If the erase block is totally defective and contains no sectors which can be written to (signified by a specific entry in the erase block table) then the erase pointer is incremented again. When the erase pointer reaches the highest usable erase block it then goes back to the lowest part of the memory and points at the lowest usable erase block. Erase blocks are erased at a rate which aims to maintain an approximately constant number of erased sector locations immediately above the current write location. An erase operation is therefore performed when the difference between the sector numbers defined in the erase pointer and write pointer falls below a certain threshold value. Prior to erasure of an erase block, a sector containing valid data within that erase block must be relocated. Relocation of this valid data is accomplished by reading the sectors containing valid data into a relocate buffer (hereinafter referred to as a transfer buffer) and writing them back to the location pointed at by the write pointer. The logical to physical address translation table is updated to point to the new physical sector address for the sector. Since only valid sectors are relocated in this way and since sectors with obsolete data are automatically erased, automatic compaction of the stored data is provided.

The process established by this algorithm can be considered equivalent to a small contiguous region of erased memory cycling through the entire FLASH memory space, removing fragmented sectors it encounters at its upper boundary and compacting and relocating them at its lower boundary.

It will be appreciated that various modifications may be made to the above described embodiments within the scope of the present invention. For example, the system would work equally well if the contiguous region of erased memory cells moved in the other direction: that

is, if the lower boundary removed fragmented sectors which were immediately below the lower boundary and valid data was written to the location immediately below the upper boundary.

## Claims

1. A solid state memory for emulating a disk drive comprising: translation means for translating a logical address to a second address for accessing a store; a store composed of non-volatile memory cells erasable in blocks; characterised in that a first pointer is used to point to an unwritten location in the store, and a second pointer is used to point to the next unerased erasable block in sequence to the erasable block containing the said unwritten memory location; control means being provided to ensure that there is always at least one erasable block in the erased condition between the first and second pointers.
2. A solid state memory according to claim 1 where the said unwritten location in the store is the lowest unwritten location in the store.
3. A solid state memory according to claim 1 where the said unwritten location in the store is the highest unwritten location in the store.
4. A solid state memory according to either of claims 1 or 2 where the said next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is the unerased erasable block with the lowest physical address above the first pointer.
5. A solid state memory according to either of claims 1 or 3 where the said next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is the unerased erasable block with the highest physical address below the first pointer.
6. A solid state memory according to any preceding claim where the said next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is determined by a suitable algorithm.
7. A solid state memory according to any preceding claim where a buffer is used to increase the speed of writing data to the the store.
8. A solid state memory according to claim 7 where the said buffer is an SRAM buffer.
9. A solid state memory according to claim 7 where the said buffer is a DRAM buffer.
10. A solid state memory according to any preceding claim where the said store is FLASH EPROM.
11. A solid state memory according to any preceding claim where the said the store is composed of chalcogenide cells.
12. A solid state memory according to any preceding claim where the said the store is a ferro-

optic memory.

13. A solid state memory according to any preceding claim where the said control means are provided by a microprocessor, a microcontroller, or a state machine.



Application No: GB 9519670.5  
Claims searched: 1-13

Examiner: B.G. Western  
Date of search: 10 November 1995

**Patents Act 1977**  
**Search Report under Section 17**

**Databases searched:**

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.N): G4A AMG1 AMX

Int Cl (Ed.6): G06F 3/00 12/00 12/02 12/06 12/08 12/12

Other: On-line : WPI, INSPEC, COMPUTER

**Documents considered to be relevant:**

Category	Identity of document and relevant passage	Relevant to claims
X	GB-2251324-A INTEL (N.b pages 30-36)	1-10,13
X	EP-0522780-A2 IBM (N.b Figure 3, columns 7-14)	1,6-10,13
A	WO-94/20906-A1 M-SYSTEMS See whole document	-

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.